UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/608,587	06/27/2003	James M. Sweet	D/A2555	8422

25453   7590   01/04/2008
PATENT DOCUMENTATION CENTER
XEROX CORPORATION
100 CLINTON AVE., SOUTH, XEROX SQUARE, 20TH FLOOR
ROCHESTER, NY 14644

EXAMINER
HILLERY, NATHAN

ART UNIT	PAPER NUMBER
2176	

MAIL DATE	DELIVERY MODE
01/04/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

	Application No.	Applicant(s)
	10/608,587	SWEET ET AL.
Office Action Summary	Examiner	Art Unit	
	Nathan Hillery	2176	

-- *The MAILING DATE of this communication appears on the cover sheet with the correspondence address* --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>11 October 2007</u>.

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) <u>1-37</u> is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) <u>1-28 and 35-37</u> is/are rejected.

7)☒ Claim(s) <u>29-34</u> is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☐ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
    Paper No(s)/Mail Date _____.

4)☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.

5)☐ Notice of Informal Patent Application (PTO-152)

6)☐ Other: _____.

DETAILED ACTION

1.   This action is responsive to communications: RCE filed on 10/11/07.

2.   Claims 1 – 37 are pending in the case. Claims 1, 16 and 27 are independent.

Continued Examination Under 37 CFR 1.114

3.   A request for continued examination under 37 CFR 1.114, including the fee set

forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this

application is eligible for continued examination under 37 CFR 1.114, and the fee set

forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action

has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on

10/11/07 has been entered.

Claim Rejections - 35 USC § 103

4.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

5.   Claims 1 – 6, 10 – 13, 16 – 20, 25 – 28, 36, and 37 are rejected under 35

U.S.C. 103(a) as being unpatentable over Bharat et al. (6112203) and further in view of

Earl (5924104).

6.   **Regarding independent claim 27**, Bharat et al. teach that we locate pages that

point to at least one of the pages in the start set 201. We call this set of pages the back

set 202 (Column 4, line 61 – Column 5, line 20), which meets the limitation of

performing a page-level link analysis that identifies those hyperlinks on a page linking to a candidate document page.

Bharat et al. teach that if a link points to a page that is represented by a node in the graph, and both pages are on different servers, then a corresponding edge 213 is added to the graph 211. Nodes representing pages on the same server are not linked. This prevents a single Web site with many self-referencing pages to unduly influence the outcome (Column 4, line 61 – Column 5, line 20), which meets the limitation of **identifying possible progression links; identifying possible table of content links;** and **examining the possible progression links and the possible table of content links for common characteristics**, since the specification states that "the possible progression links 230 and possible table of content links 240 are passed to module 250 for a final examination to weed out links which have properties that are not characteristic of typical intra-document links, e.g. they point to a different web server" (p 7, lines 26 – 30). It should be noted that pages on the same server are nodes and are thus still apart of the resulting graph.

Furthermore, it should be noted that the self-referencing pages of Bharat et al. are equivalent to intra-document links and that those intra-document links can be "possible" progression and/or table of contents links, since the Office has interpreted the word "possible" as "could be" and within the broadest, reasonable interpretation in light of the specification, which states that a link analysis phase consists of the identification for a given hypertext page of the most likely desirable intra-document links. Those intra-document links fall into two categories: progression links and table of contents links (p

5, second paragraph). Thus, any intra-document link – a link that points to a different web server – could be a possible progression or table of contents link.

Bharat et al. teach that a larger n-graph 211 can be constructed by repeating this process for the back and forward sets 202-203 to add more indirectly linked pages (Column 4, line 61 – Column 5, line 20), which meets the limitation of **performing a recursive application of the page-level link analysis to the linked candidate document page and any further nested candidate document pages thereby identified, until a collective set of identified candidate document pages is assembled.**

Bharat et al. do not explicitly teach **performing a document-level analysis that examines the collective set of identified candidate document pages for grouping into one or more documents; examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of typical intra-document links, to provide a resultant set of identified candidate document pages; and grouping the resultant set of candidate document pages by an automated into a document representation stored in memory by the automated system; and printing or viewing on a display by a user the document representation.**

Earl teaches that the link display manager 300 includes a display system for defining predetermined screen element properties providing visual cues for distinguishing the identified links 202 and 204. When a user provides an input link selection to select a new document, the document parser 304 parses the selected new

document to identify intradocument links 202 and interdocument links 204 (Column 2, line 59 – Column 3, line 9), which meets the limitation of **examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages**.

Earl teaches that the display system 306 processes the identified intradocument links 202 and interdocument links 204 for displaying distinctively the intradocument links 202 and interdocument links 204 with predetermined visual cues to differentiate the links 202, 204 (Column 2, line 59 – Column 3, line 9), which meets the limitation of **grouping the resultant set of candidate document pages by an automated into a document representation stored in memory by the automated system; and printing or viewing on a display by a user the document representation**.

It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the invention of Bharat et al. with that of Earl because such a combination would provide the users of Bharat et al. with *an improved method and apparatus for displaying links on a user display interface in a computer system* (Column 1, lines 39 – 41).

7.   **Regarding dependent claims 3 – 6, 17 – 20, and 28,** Bharat et al. teach that the nodes in the start set are first scored according to their connectivity, and the number of terms of the query that appear as unique sub-strings in the URL of the represented documents. The score is a weighted sum of the number of directed edges to and from a node and the number of unique sub-strings of the URL that match a query term

(Column 3, lines 10 – 15), which meet the limitation of **the page-level link analysis includes examination of contextual clues, the contextual clue is a particular class of content item associated with the hyperlink, the class of content item is a class of text, the class of text is a directional word or phrase**.

8..   **Regarding dependent claims 36 and 37**, Bharat et al. teach that we assign a similarity weight to each node 213 of the sub-graph 255. Various document similarity measuring techniques have been developed in Information Retrieval to determine the goodness of fit between a "target" document and a collection of documents. These techniques typically measure a similarity score based on word frequencies in the collection and a target document (Column 6, lines 51 – 57), which meet the limitation of **the page-level analysis includes determining the similarity of the hyperlink destination to that of other hyperlinks within the page, and the page-level analysis includes determining the similarity of the hyperlink destination to the location of the current page.**

9.   **Regarding claims 1, 2 and 10 – 13**, the claims incorporate substantially similar subject matter as claims 27, 36 and 37 and are rejected along the same rationale.

10.   **Regarding claims 16, 25 and 26**, the claims incorporate substantially similar subject matter as claims 27, 36 and 37 and are rejected along the same rationale.

11.   Claims 7 – 9, 14, 15, 21 – 24, and 35 are rejected under 35 U.S.C. 103(a) as

being unpatentable over Bharat et al. (US 6112203 A) and Earl (5924104) as applied to

claims 1 – 6, 10 – 13, 16 – 20, 25 – 28, 36, and 37 above, and further in view of Prince

(US 6877002 B2).

12.   **Regarding dependent claims 7 – 9, 14, 21 – 24, and 35**, neither Bharat et al.

nor Earl explicitly teach **the class of content item is a class of image, the class of**

image is an image containing a directional symbol, a textual clue is obtained for

the image, the identifying of table of content links includes the presence of at

least one other hyperlink nearby with the page description.

        However, Prince teaches that the parsed results (from step 42 in FIG. 4) relating

to the media are passed to extraction agent 68 via an extraction queue 67. The

extraction queue 67 comprises URLs to be analyzed with respect to associated media

metadata. The extraction queue 67 may comprise metadata queue entries such as

media URLs, Web page URLs, Web page titles, Web page keywords, Web page

descriptions, media title, media author, and media genre. Each queue entry added to

the extraction queue is assigned a processing time and a priority (Column 7, lines 23 –

37), meets the limitation of **the class of content item is a class of image, the class of**

image is an image containing a directional symbol, a textual clue is obtained for

the image, the identifying of table of content links includes the presence of at

least one other hyperlink nearby with the page description.

        It would have been obvious to one of ordinary skill in the art at the time of the

invention to combine the combined invention of Bharat et al. and Earl with that of Prince

because such a combination would allow the users of Bharat et al. and Earl the benefit

of *A method for querying metadata associated with media on a computer network*

includes separating the metadata into keywords (Column 2, lines 37 – 39).

13.   **Regarding claim 15**, Bharat et al. teach that we assign a similarity weight to

each node 213 of the sub-graph 255. Various document similarity measuring

techniques have been developed in Information Retrieval to determine the goodness of

fit between a "target" document and a collection of documents. These techniques

typically measure a similarity score based on word frequencies in the collection and a

target document (Column 6, lines 51 – 57), which meet the limitation of **the page-level**

analysis includes determining the similarity of the hyperlink destination to that of

other hyperlinks within the page.

Allowable Subject Matter

14.   Claims 29 – 34 are objected to as being dependent upon a rejected base claim,

but would be allowable if rewritten in independent form including all of the limitations of

the base claim and any intervening claims.

Response to Arguments

15.   Applicant's arguments filed 10/11/07 have been fully considered but they are not

persuasive. .

16.   Applicant argues that Earl does not teach **examining the collective set of**

identified candidate document pages to weed out links which have properties that

are not characteristic of intra-document links, to provide a resultant set of

identified candidate document pages because what Earl defines as intra-document is what Applicant would call intra-page and thus what Earl calls inter-document is really inter-page (pp 12 and 13).

The Office disagrees.

In light of all the evidence, the claim recites **links, which have properties that are _not characteristic_ of intra-document links**. Consequently, the Office maintains that Earl clearly and explicitly teaches intra-document and inter-document links that meet the claimed intra-document link, since both or at least one of the links, intra-document and inter-document, have properties that are characteristic of intra-document links within the broadest, reasonable interpretation in light of the Specification.

17.   Applicant argues that Earl further does not teach **examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages** because Earl keeps all the links while applicant's invention discards or "weeds" out the undesirable links (pp 14 and 15).

The Office disagrees.

First, Applicant is correct that Encarta defines "weed out" as "to separate out something undesirable. However, the Office maintains that Earl does "weed out" the links within the broadest, reasonable interpretation in light of the specification. Thus, Earl, by applicant's own admission, teaches discriminating visually between intra-document and inter-document links, which meet the definition of separating out, or

weeding out, the links visually on screen. The requirement to have to discard the links is still too limiting in view of what is actually claimed and in light of all the evidence.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Nathan Hillery whose telephone number is (571) 272-4091. The examiner can normally be reached on M - F, 10:30 a.m. - 7:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doug Hutton can be reached on (571) 272-4137. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Nathan Hillery
Examiner
Art Unit 2176

NH